HumRRO

Interim Evaluation Report
IR–01–29r

June 29, 2001

California High School Exit Examination (CAHSEE): Year 2 Evaluation Report

Lauress L. Wise
D.E. (Sunny) Sipes
Carol E. George
J. Patrick Ford
Carolyn DeMeyer Harris

Prepared for:
California Department of Education
Sacramento, CA

Contract Number: 9234

Human Resources Research Organization
66 Canal Center Plaza, Suite 400 • Alexandria, VA 22314-1591

California High School Exit Examination (CAHSEE): Year 2 Evaluation Report

Executive Summary

Background

California has moved through the second year of its schedule for requiring graduation exams in mathematics and ELA beginning with the Class of 2004. As is the case in nearly half of the states in the country, California began this initiative in response to widespread support for high standards and for some mechanism that holds students to them. This component of California's testing program is intended to ensure that all students graduating from high school can demonstrate grade level competency in reading, writing, and mathematics. The California Education Code, Chapter 8, Section 60850, specifies requirements for the California High School Exit Examination (CAHSEE). Since January 2000, the California Department of Education (CDE) has worked with a development contractor, the American Institutes for Research (AIR), throughout the development and tryout of test items for use in the CAHSEE and to develop and implement procedures for operational administration, scoring, and reporting. The first operational administration to 9th graders on a voluntary basis was completed in March and May of 2001. Results from these administrations will be released in August 2001.

The California legislation specifying the requirements for the new exam also called for an independent evaluation of the CAHSEE. CDE awarded a contract for this evaluation to the Human Resources Research Organization (HumRRO). HumRRO's efforts focus on analyses of data from the field test of items (test questions), the field administration of the test, the annual administrations of the CAHSEE, and use of these analyses to report on trends in pupil performance and pupil retention, graduation, drop-out, and college attendance rates. As specified in the legislation, the evaluation reporting will include recommendations for improving the quality, fairness, validity, and reliability of the examination. This report describes evaluation activities through June 2001, summarizes the results of these activities, and offers initial recommendations based on conclusions drawn from these results. It should be noted that this is a report of yearly activities; we have had a relatively short time frame in which to examine the operational test and longitudinal survey results. The current report is a contractual requirement and not one of the reports mandated in the legislation specifying the evaluation. More comprehensive results from the March and May administrations will be included in the next mandated evaluation report required by February 2002.

There were four main activities in Year 2 of the evaluation:

- HumRRO conducted a special census survey of all high school districts in California at the request of the State Board of Education (SBE). Results from this survey, which examined awareness of CAHSEE, preparations and expectations for the exam, and baseline student outcomes, were reported fully at the end of the fall semester (Sipes, Harris, Wise, & Gribben, 2001).
- Researchers analyzed data from the Fall 2000 Field Test of CAHSEE questions. Results of these analyses are presented in Chapter 2 of this report.

- HumRRO personnel observed the March and May 2001 operational administration of the CAHSEE, analyzed the results available from the March administration, and reviewed plans for reporting, including determination of the minimum passing scores. Results of these activities are described in Chapter 3 of this report.
- The research team conducted a Spring 2001 survey of teachers, principals, and test coordinators in the longitudinal sample of schools we are following as part of our evaluation. Results from the test coordinator survey are included in Chapter 3. Results from the teacher and principal surveys are presented in Chapter 4 of this report.

Summary of Year 2 Activities and Results

CAHSEE Fall 2000 Field Test. Results of the Spring 2000 Field Test indicated that nearly all of the items had acceptable statistical properties and could be used on operational CAHSEE forms. Additional test questions, however, were needed to cover particular standards and to support the assembly of multiple test forms. Additional test questions were developed by AIR and included in a second field test conducted in Fall 2000.

HumRRO's analyses address the following three general issues:

- What proportion of items has good statistical properties?
- Were the questions included in the second field test significantly different in quality and difficulty from the questions in the first field test?
- How difficult are the questions that address specific standards and did the difficulty level vary among different demographic groups?

The test questions in the Fall Field Test were found to be of similar difficulty and quality in terms of statistical properties to the questions in the Spring 2000 Field Test despite the fact that the Fall Field Test questions were newly developed and had not been subjected to extensive prior screening. For each subject, 20 questions from the Spring Field Test were repeated in each of the Fall Field Test forms to provide a means for adjusting item difficulties for differences between the two field tests in student achievement levels. Tenth graders in the Fall Field Test performed somewhat worse (a drop of 4.5 in the average percent correct responses) on math questions in comparison to 10^{th} graders in the Spring Field Test who had had seven more months of instruction. For the ELA questions, however, students in the Fall Field Test performed slightly better. Analyses of the questions by content standard indicated that there were sufficient questions for each standard to construct several unique test forms. The relative difficulties of questions for different standards were similar to those reported in our Supplemental Year 1 Report (Wise, Sipes, Harris, Collins, Hoffman, & Ford, 2000).

Observation and Analysis of the March 2001 Operational Administration. Chapter 3 presents our observation and analyses of the results of the March 2001 administration of the CAHSEE. The first section of this chapter describes test administration issues. HumRRO observed focus groups of district testing coordinators, a training workshop for test coordinators, and administration of the CAHSEE at three sites. In addition, a survey was administered to test coordinators at the schools in our longitudinal sample (described in Chapter 4). Findings indicated that while the schools varied in the ways they conducted CAHSEE, school staffs were well prepared and generally provided good test conditions. The

most striking overall feature was how seriously the students took the test. Logistical issues at school sites included balancing extended time with test security, particularly for the ELA exam. One other issue was that both our observations and our survey indicated a low frequency of use of testing accommodations.

In examining results from the March administration, HumRRO staff computed item statistics and found that items performed close to original expectations with respect to the difficulty and information value of each item. Staff observed item-scoring procedures for the two essay questions and analyzed the consistency of scoring results. Two different readers judged each essay and sufficient agreement was reached more than 99% of the time for the first essay and roughly 98% of the time for the second essay. Where disagreements did occur, there was a systematic process for their resolution.

HumRRO examined the process for setting minimum passing scores. The standards-setting process included a reasonable mix of teachers, other educators, parents, and businessmen and women who were broadly representative of their peers across the state. The standards-setting process was well specified and engendered a relatively deep discussion of the skill requirements of specific items and the importance of these requirements. Some panel members were surprised at the relatively low passing rates for the standards they had proposed. Following discussions, few wanted to change the standards and, in the end, the median ratings did not change. Both the mathematics and ELA panels recommended that the minimum passing score be set at about 70 percent of the total possible points on each test. The SBE subsequently concurred with a recommendation from the Superintendent to adopt initial passing criteria that were more lenient. The passing criteria for the Class of 2004 were set in recognition of the fact that the new content standards were not yet in place when these students were in earlier grades where essential prerequisite skills are taught. The passing levels approved by the Board, 60% of the possible points for ELA and 55% for math, are provisional pending review of results for 10^{th} graders next year.

Using the passing levels set by the Board at its June 2001 meeting, we examined passing rates for students who participated in the March administration. Overall, 65% of the students tested in March passed the ELA exam and 45% passed the math exam. Passing rates for students with disabilities in the March administration were considerably lower, at 22% for E-LA and 12% for math. Not surprisingly, passing rates for math varied systematically by the pattern of math courses completed or in progress, ranging from a passing rate of over 90% for students who had completed algebra 1 and were currently enrolled in geometry, down to 18% for students who had not taken and were not currently enrolled in algebra 1. In schools where 500 or more students were tested, passing rates ranged from below 10% to above 90%.

At the end of the CAHSEE exams, students completed a brief questionnaire on their reactions to the test and their plans for high school and beyond. HumRRO examined the responses to these questions separately for students who did or did not pass each of the two tests.

Our analyses of results from the March administration also included an assessment of the accuracy of pass/fail classifications. Based on statistical estimates of measurement error, we defined a “zone of uncertainty” where students were close enough to the minimum passing score that there was some potential for classification to be affected by measurement error.

Inside this zone of uncertainty (defined as the range of scores for which the probability of classification error exceeded 10%) about 70% of the students were correctly classified as passing or failing the test; outside this zone 98% or more of the students were correctly classified. For math the zone of uncertainty was relatively narrow—only 6 score points. Only 12% of the students tested were within this "too close to call" range. For ELA, the zone was a bit wider, 13 of the 90 possible score points, and contained about 20% of the students tested. While the level of uncertainty may seem high, it is no greater than with other testing programs with which we are familiar. In fact, for examinees near the borderline, there will always be some uncertainty, but the consequences of incorrect classification decisions in these cases are not great, particularly where retesting is allowed.

Spring 2001 Survey of Teachers and Principals. Chapter 4 describes results from the second spring survey of teachers and principals from our longitudinal study sample. Issues focused on *awareness, planning and preparation, alignment, expectations, and potential outcomes*. Surveys were administered following the Spring 2001 CAHSEE administrations but prior to results being provided to the schools. Survey results indicated that, overall, both principals' and teachers' awareness of the CAHSEE (knowledge of skills covered and familiarity with administration plans) increased from last year. Similarly, principals' ratings of student and parent familiarity with CAHSEE increased from last year.

With respect to alignment, responses indicated a slight increase in estimated preparedness of students in 9th grade from 2000 to 2001 and a larger increase in predicted preparedness of students in 10th grade. Teachers were asked to identify courses in which particular standards were taught. Many of the courses identified are typically taken during the 10th grade, reinforcing the idea of deferring initial testing until 10th grade.

HumRRO assessed the potential consequences of CAHSEE by examining predicted pass rates, impact on student motivation and parental involvement, and impact on instructional practices. Predicted pass rates, collected before the discussion of passing levels by the State Board, were similar to last year's predictions and, on average, were reasonably comparable to actual results. A slightly more positive impact on student motivation and parental involvement was predicted for students and parents prior to the first administration than upon receiving pass/fail results from the first attempt. Predictions of the impact of the CAHSEE on student retention and drop-out rates were generally similar in 2000 and 2001, although principals' predicted impact on student drop-out rates were slightly more negative this year. Teachers continue to expect the CAHSEE to have a positive impact on instruction, and they generally expect that impact to grow increasingly positive over time.

Principals were asked to indicate what actions the school plans to take or has implemented to promote learning for all students. Responses indicate that while a number of actions have already been undertaken to promote student learning, many of these actions have only been partially implemented at this time.

Key Findings and Recommendations

In our earlier evaluation reports, we expressed concern with the time line for implementing the new graduation requirement. Our concern was based on two key questions:

(1) Would the exam be ready for the students?

(2) Would students be ready for the exam?

The first question was asked with regard to the risk of problems in the assembling and printing of test forms, with the administration of the test, and with the reporting of results. Based on evaluation activities to date, we offer the following general findings:

General Finding 1: Progress in developing the exam has been noteworthy. We found no significant problems with the exam administered in March 2001 or with plans to report results from that administration.

Given low initial passing rates, there may be a tendency to question the validity of the exam. Our analyses of data from the March 2001 administration, however, showed that all test questions performed as expected. Forms were printed correctly and on time and delivered to districts with few difficulties. Administration of the exam presented a number of significant challenges to schools in finding times and spaces in which to schedule students to take the exam. Even though the March administration was not a practice test, as it appeared for awhile it that might be, it provided a good opportunity to identify logistical and administrative issues to be addressed further in future administrations. The 2002 administrations will be the first time students who have completed much of the 10^{th} grade curriculum will take the exam. Lessons learned from the 2001 administrations should be helpful in improving the process for 2002.

General Finding 2: The process used to establish minimum passing scores was well designed and executed and the resulting passing standards appear reasonable.

There was some concern that the passing scores for the two exams could not be set until data from a census testing of 10^{th} graders were available. With the failure of the urgency legislation (SB 84), SBE was required to set minimum passing scores without normative information on 10^{th} graders. Many experts disagree with the use of normative information and, where it is used, it rarely has much impact on the recommendations of the standards-setting process. CDE and AIR used a systematic process for identifying panels of teachers and others who were very familiar with California standards and students and were broadly representative of the state. The SBE appropriately considered the passing standards as provisional, recognizing concerns that results for students completing the 10^{th} grade curriculum are not yet available.

General Finding 3: Progress on providing all students adequate opportunity to learn the material covered by CAHSEE has been good, but it is too soon to tell whether there will be significant problems in preparing students in the Class of 2004 to pass the exam.

Since our earlier reports expressed concern as to whether all schools could provide the Class of 2004 adequate opportunity to master the standards tested by CAHSEE, a number of changes have occurred:

1. Beginning with the Class of 2004, algebra will be a statewide requirement for high school graduation.

2. Survey results indicate that schools are taking the content standards seriously and have progressed in plans to provide students opportunities to learn these standards.

3. Principals and teachers report that students and parents have a greater awareness of CAHSEE than they did a year ago.

4. SBE plans are in place for adoption of K-8 textbooks aligned to the content standards and to incorporate results of standards-based tests into the Academic Performance Index (API).

5. CDE has launched a campaign for disseminating information about the CAHSEE and the content standards that it covers to districts and schools.

The fact that significant numbers of 9th graders have not yet mastered the standards covered by CAHSEE is not surprising. Results from our Spring 2001 survey suggest that many of the standards are covered by courses most students do not take until the 10th grade. Members of the standards-setting panels were generally optimistic about schools' capacity for bringing students up to standard.

General Recommendation 1: Stay the course. The legislature and Board should continue to require students in the Class of 2004 to pass the exam, but monitor schools' progress in helping most or all of their students to master the required standards.

Notwithstanding earlier recommendations, we think it best not to alter the current schedule for implementing the CAHSEE requirements at this time. As expected, initial passing rates are low, indicating that many 9th grade students have not yet had the opportunity to learn the material covered by the CAHSEE. Continuing with the current requirement means demanding that schools, teachers, and even parents not give up on the Class of 2004 just because their education to this point may not have been as comprehensive as we would like it to be. Most educators with whom we have spoken are optimistic regarding the potential for most students to master the required content standards given more years of instruction and targeted assistance. Schools and districts have expended considerable effort in improving the curriculum to increase coverage of the state content standards, particularly those covered by CAHSEE. A decision to delay the requirement at this point could be seen as undercutting these efforts.

While we think the state should continue to move ahead, we continue to have concerns, as expressed in our earlier reports and reflected in current discussion over Assembly Bill AB-1609 as to whether all students in the Class of 2004 will have adequate opportunity to learn

the material covered by the CAHSEE by the time they complete the 12th grade. Evidence of opportunities to learn, based on analysis of the curriculum, is, as suggested by some, necessarily limited. ***However, the best evidence that a school system is providing its students adequate opportunity to learn the required material is whether most students do, in fact, learn the material.*** Our evaluation will continue to monitor passing rates by school as an indicator of the extent to which students in these schools have had effective opportunities to learn the required knowledge and skills. A critical factor will be whether schools with the most difficult challenges, as evidenced by initial passing rates, will be given the guidance and resources needed to bring their students up to required levels.

Whether the requirement is deferred or not, it will be very important to give the CAHSEE requirement time to work. The history of state assessment programs shows a lack of stability over any prolonged period of time. For students to achieve the skills embedded in California's content standards, success may take a sustained effort over an extended period of time. "Staying the course" will be required to allow this to happen.

General Recommendation 2: The legislature and Board should continue to consider options for students with disabilities and English learners.

There is significant tension between the desire to have high expectations for all students, including students with disabilities and English learners, and the need to be realistic about what some students can accomplish. Initial low passing rates for both of these groups suggest particular concern with the time it may take to help these students master the required standards. Options to be considered range from more liberal use of accommodations, to some form of alternative diploma for students who cannot reasonably be expected to develop or demonstrate the required skills, and also to deferring the graduation requirement for these students.

Other Specific Findings and Recommendations

Our Year 2 Evaluation Report contains a number of other, more specific findings and recommendations. These include:

1. **More technical oversight is needed.** Because of the rapid pace of implementation, a number of decisions have been made without technical review of the consequences. Examples are the decision to shorten the tests without public consideration of consequences for test score accuracy and the lack of review of plans for equating scores from the different test forms used in March and May.

2. **For future classes, testing should be delayed until the 10th grade.** Many students do not receive instruction in important content standards until the 10th grade. Other options should be available for assessing the readiness of 9th graders to pass this exam.

3. **A practice test of released CAHSEE items should be constructed and given to districts and schools to use with 9th graders to identify students at risk of failing the CAHSEE.** Scoring instructions should be included so that teachers and students can gauge how much additional effort might be needed to reach passing levels. The

practice test should include as much diagnostic information as possible. Alternatively or in addition, research showing linkage between the 8th and 9th grade California Standards Test used for school accountability would support use of scores from this assessment to identify students who need additional help to pass the CAHSEE.

4. **More extensive monitoring of test administration and a system for identifying and resolving issues is needed.** Observation of the initial administration revealed some concern about describing and enforcing procedures for test session breaks so as to maintain test security. In addition, procedures for determining appropriate testing accommodations may need further clarification and reinforcement.

5. **The state needs a more comprehensive information system that will allow it to monitor individual student progress.** It is not clear that school information systems will necessarily support passing along information on problems associated with transfer students who have passed or not passed part or all of the CAHSEE. In addition, research databases on cumulative passing rates for each high school class and on the relationship of CAHSEE scores to results from other tests are needed to answer important policy questions. A mechanism for creating such databases without infringing on student privacy concerns is needed.

6. **The legislature should specify in more detail how students in special circumstances will be treated by the CAHSEE requirements.** A number of students may not have the full range of opportunities to take the CAHSEE. These include students who transfer into the state in the 12th grade, students in the Class of 2003 who, through illness or other unforeseen circumstance, fail to graduate on time and will then be subjected to requirements for the Class of 2004, and English learners who may be exempted from taking the CAHSEE until late in their high school years. Such students would miss out on several opportunities to pass the CAHSEE and end up with at most 3 or 4 chances to pass the test rather than the 8 chances most students would have.

More detailed explanations and rationales for each of these recommendations are presented in the full text of the report.

California High School Exit Examination (CAHSEE):
Year 2 Evaluation Report
Table of Contents

List of Tables

List of Figures